Alternative Solutions of the Paradoxes?

- Paraconsistency seems to be a desperate measure. If there is any other way out of the antinomies ("paradoxes") one may rather go that way.
- In this chapter we look at some suggestions for some antinomies.
- Whereas paraconsistency provides a *universal* method to deal with any antinomy whatsoever, or even some proponents proclaim a universal logic providing the framework for all antinomy related reasoning the alternatives are a rag back of solutions many of which deal only with just one antinomy. Nevertheless the case by case comparison may be fruitful given the philosophical worries accompanying paraconsistency.
Russell‘s Antinomy and Lesniewski

- Russell‘s antinomy (usually called "Russell‘s Paradox") was one of the main motivations for Lesniewski to develop his mereology.
- Mereology is the theory of parts and wholes. It substitutes the concept of an aggregate for the concept of a set. Usually mereology is pursued for ontological reasons within a nominalist programme. Lesniewski, however, saw it also as superior to set theory in its avoidance of the antinomies.
- We look at some version of Lesniewski's mereology (ideas from 1913-1914) in which he avoids the Russell Paradox by having all classes (i.e. aggregates) being members of themselves. Further on in this version of mereology there is no empty class. So the Russell Set – or the corresponding mereological class – looking for the classes that are not members of themselves just does not exist!
Lesniewski on Russell's Paradox

- Lesniewski's mereology is built on a few definitions and axioms:
  - (1) x is element of the class y if there is a z such that
    (i) y is the class of objects z
    (ii) x is a z.
    formally: \((\forall x,y)(x \in el(y) \equiv (\exists z)(y \in Cl(z) \land x \in z))\)
  This is the definition of being an element. Lesniewski uses "\(\in\)" as elementary predication/being-such relation. [Note also that by merely being a part of an aggregate the part has not a defining property, therefore this twofold definition.]
- Classes are objects: (2) If y is a class, then y is an object.
  formally: \((\forall y)(y \in C \supset y \in y)\)
  So "\(\in\)" has ontological impact.
- If something is an object, it is its own class:
  (3) If x is an object, then x is the class of objects x.
  formally: \((\forall x)(x \in x \supset x \in Cl(x))\)
  So an object also defines its own kind, so to say.
Lesniewski on Russell's Paradox (II)

- From (2) and (3) we get:
  (4) If \( y \) is a class, then \( y \) is a class of objects \( y \), and \( y \) is an element of \( y \).
- By generalizing on the second "\( y \)"s:
  (5) If \( y \) is a class, then there is a \( z \), such that \( y \) is the class of objects \( z \) and \( y \) is an element of \( z \).
- From (5) and (1) taking "\( x \)" in (1) to be "\( y \)":
  (6) If \( y \) is a class, then \( y \) is an element of the class \( y \). Thus we have established that all classes are self-members!
- Mereology takes classes as existing aggregates/heaps. As an axiom we have that the aggregation of objects is *not nothing*, i.e. the class cannot be empty:
  (7) If some object is the class of objects \( y \), then some object is \( y \).
  [Thus in this version of mereology there is no empty set.]
- As a dual of (6) we get:
  (8) No object is a class which is not element of itself.
  [Note that, by (7), if \( y \) is not element of \( x \), still both are objects, so the negation of a statement of elementhood cannot mean that the class \( x \) doesn't exist.]
Lesniewski on Russell's Paradox (III)

- And now from (7) and (8) we derive the non-existence of the Russell Set:
  
  \[(9) \text{ No object is a class of classes which are not elements of themselves.}\]

  We argue by (8) that there are no objects corresponding to the definition of the Russell Set, and by (7) that there is no aggregate of nothing.

- Lesniewski's later system (1916) – as mereologies today – focuses more on the relation of being a part. In that system no object is a part of itself (being a part is defined as asymmetrical), but still all classes are members of themselves, there is no empty class, and being a member of $x$ is equivalent to being a subset of $x$. That system still remains in a countable domain.
Assessment of the Mereological Solution

- Lesniewski's Mereology can avoid Russell's Paradox. Some other antinomies do not occur in Lesniewski's mereology because of its limitations of cardinality.
- The main problem with mereology is whether it is a revision of standard mathematics, by giving up its foundation (standard set theory) in favour of mereology. Lesniewski's mereology is such a revision. Lewis' mereology is not; it is just a recasting/reduction of standard set theory as a *façon de parler* founded in mereological axioms. Lewis achieves (with the residual problem of singletons) the nominalistic reduction of set theory to heaps. Lewis also provides the higher cardinalities. Since it is a reduction, however, Lewis has still all the problems we had with standard set theory (e.g. absence of a universal set). A revision of standard mathematics is not so easily advertised to mathematicians, who usually don't care about the philosophical problems given with ZFC or ontology. The problem of an adherent of a Lesniewski like mereology, therefore, is to regain standard mathematics. (The fundamental idea of Lewis was to avoid just this.)
Set Theories with Universal Sets

- One philosophical problem with standard set theory besides restricted Comprehension is the absence of a universal set (as the extension of the concept set). Some set theories with universal sets have been proposed.
- The widest known set theory with a universal set $V$ – for which we have $V \in V$ – is Quine's NF (from his paper "New Foundations of Mathematics", 1937). This system has been explored in its relation to type theory and ZFC. Another such set theory is (Church 1974).
- Like dialetheist set theories, which work "only" by Extensionality and Naive Comprehension, NF works by Extensionality and an Comprehension Scheme that is not as restricted as in ZFC.
- Quine's Comprehension Scheme uses the idea of stratified formula (similar to the theory of types):
  $$ (\exists y)(\forall x)(x \in y \equiv P(x)) $$
where "$y$" is not free in "$P(\ )$" and "$P(\ )$" is stratified. (A formula is stratified if the set on the right hand side of "$\in$" is of a higher level than that on the left, and its definition does not include that on the left.)
Set Theories with Universal Sets (II)

- **NF** has the power to introduce Pairing, (Absolute) Complement, Powerset, Union as instances of Comprehension.
- Non-stratified formula can be used in **NF** (this is different to Type Theory), but they cannot be used to define sets. Since non-stratified formulas can be used in **NF** one does not need a universal set or an empty set for every level (as in Russell's Type Theory) to have well-formed formula.
- The antinomies – especially the Russell Set – are avoided, since the corresponding open formulas in the Comprehension Scheme are not stratified.
- $V \in V$ means that Cantor's Theorem does not hold (in general) in **NF**; but the set of units sets of its elements is smaller than $V$ itself!
- **NF** has other highly controversial features like the existence of infinite descending chains of cardinals, what conflicts with the Axiom of Choice (which thus doesn't hold in **NF**). Some functions (like the successor function) are *not* part of the universe – so where are they?
Set Theories with Universal Sets (III)

- A version of \textbf{NF} that tries to avoid many of the peculiarities of \textbf{NF} is \textbf{NFU} (\textbf{NF} with urelements), developed by Randall Holmes.
- \textbf{NFU} is consistent, it is consistent with the Axiom of Choice!
- \textbf{NFU} is an extremely strong set theory. It can provide models for \textbf{ZFC} by having \textit{very} large cardinals (strong inaccessibles).
- By working with a longer \textit{(finite)} list of simple axioms [like Extensionality, Complements, Unions, Singletons, Cartesian Products, Converse, Domains, Projections, Singleton Image of a Set, Choice (!), Infinity…] Stratified Comprehension can be proven as a theorem!
- The universal set is provided by its own axiom: \{x \mid x = x\} exists. The universe \( V \) can be well-ordered (equivalent to the Axiom of Choice).
- Urelements and ordered pairs (!) are taken as primitive.
- \{<x,y> \mid x \in y\} does not exist, the complement leading to antinomies.
- The large sets (like the set of ordinals or \( V \)) in \textbf{NFU} have the strange property of not being equinumerous to their singleton images! How can that be?
Set Theories with Universal Sets (IV)

- Although every object has a singleton (by axiom), the *function* giving the singleton to every object does not exist!
- Not every supposed set exists, e.g. the set of Cantorian ordinals.
- Especially problematic is that some version of *Cantor’s Theorem* is provable. A corollary is *that the cardinality of the set of singletons of of V is less than the cardinality of V.*
- This is related to *Specker’s Theorem:*

\[ |\wp(V)| < |V| \]

which is read as proof that there are (many, many) atoms besides sets in V. In fact *most* objects in V have to be atoms or pairs.

[Unlike some version of NF in (Quine 1963) atoms are not identified with their singletons. Pairs are also objects in their own right besides sets. *Specker's Theorem* asserts that many entities in V are not in subsets of V.]
Problems with a Universal Set

- Although one may argue that a universal set is one of the hallmarks of an acceptable set theory, looking at the set theories with a universal set may give one second thoughts.
- Set theories with a universal set attack the 2nd objective to be achieved by a naïve (dialetheist) set theory, having a universal set. (The 1st objective being having Naïve Comprehension.) They can be taken as the claim that a universal set can be had without dialetheism. The more general reply to the dialetheist’s claim might be that one sees more problems than benefits related to a universal set in the first place.
- The first category of objections may be called negative in as much as they argue against the supposed gains of a universal set:

(i) Since meaning isn’t extensional, fixing the meaning of “set” does not require an extension to the predicate “is a set”. The meaning of “is a set” is fixed by the axioms governing our identification of sets.

(ii) If the extension of a predicate P cannot be a set this doesn’t mean that there is no referent of P at all, it just has to be fixed singularly.
Problems with a Universal Set (II)

(iii) If consequence is spelled out model theoretically one may quantify over all appropriate structures’ equivalence classes (in terms of isomorphy); it is not obvious that some structure is left out (this would have to be one with no isomorphic structures so far); even collecting all sets into a single set looking at this set as universe leaves one with a singleton structure isomorphic to a simple singleton: compare \( V, \{ V \}, \{ \{ V \} \} \) with \( \emptyset, \{ \emptyset \}, \{ \{ \emptyset \} \} \).

- The second category of objections may be called positive as they argue directly against a universal set as violating our intuitions itself:
  (iv) As the absence of \( V \) seems to miss having a set corresponding to sethood, NF and relatives have no set \( [ \in ] \), \( \{ < x, y > | x \in y \} \), where \( [ \in ] \) is as intuitive as \( V \) is.
  (v) Like “\( \in \)“ “\( \subseteq \)” could be taken as the basic concept of set theory. Then the powerset axiom is even more fundamental than it is intuitive already. If we get Cantor’s Theorem, \( \mathcal{P}(V) \subseteq V \) can no longer be a claim about cardinalities (just like with the natural and the even numbers), but has to claim that all subsets of \( V \) are already in \( V \). This is counterintuitive, and to block Cantor’s Theorem some subsets have to be taken as non-existent (e.g. the diagonal set).
Problems with a Universal Set (III)

NFU “solves” this problem by having more atoms/pairs than sets, but this – looking at the cardinalities involved – is obscure, because this means just too many atoms or not every atom having a singleton!

• The third category of objections asks us to reconsider some ontological alternative:

  (vi) Suppose there is no set of all sets but a realm set theory is talking of. There is no need to take it as a set; supposedly everything useful can be said talking of sets only. There may be subparts of this realm (e.g. the part where the ordinals are), but again this realm/part-relation may be something like mereological composition. There is no need to have a theory of proper classes (as a theory similar to set theory).

  (vii) Once we allow for more than finite sets not every condition defines a set – remember that Naïve Comprehension is consistent on the finite sets! The limits of restricted Comprehension (like in ZFC) are just a “failure” to commit us to even more infinite sets. That doesn’t seem so bad. Given the set of finite sets, Naïve Comprehension and (restricted) Comprehension come to the same thing.
Second Order Logic

- If the main objective for an acceptable set theory is to have Naïve Comprehension then Second Order Logic can provide this without being in any way non-standard.
- Even before considering any other specific set theoretic axioms the mere calculus of Second Order Logic itself may contain a scheme for Comprehension:
  \[ (\exists X)(\forall y)(Xy \equiv \varphi(y)) \]  
  \[ X \text{ not occurring in } \varphi \]
- This scheme does not yield antinomies, since the property/set in the range of the second order variable "X" cannot be substituted for the individual variable "y".
- The scheme thus ensures that every formula of the language defines a property/set, just as it should be.
- In effect this comes down to a class theory if one uses Second Order Logic for set theory itself (using Second Order ZFC). The variable "X" etc. then range over all the subsets of the domain (i.e. over all subsets of the iterative hierarchy), thus having proper classes as denotations.
Infinitely Valued Set Theories

- Lukaciewicz introduced not only a 3-valued logic, but an infinitely valued logic where truth values are the reals in the interval [0,1].
- One can add a truth predicate to a formalization of arithmetics that uses this logic $L^\omega$ (cf. Restall 1992), the resulting system is however $\omega$-inconsistent (there being by diagonalisation a sentence that not all sentence of some kind are true, although we cannot come up with the wrong instance).
- Similarly one can add a naive Comprehension axiom to $L^\omega$ (cf. White 1979), but the logic requires an inference rule with infinite premises then.
- This means that $L^\omega$ is too strong if we want to do Naive Truth or Set Theory, where $L^\omega$ is missing *Modus Ponens* already!
Assessing Alternative Set Theories

- **L**^κ for the properties just mentioned is no option for the dialetheist.
- **NF, NFU** and related systems come with a universal set, but have highly counter-intuitive properties/meta-theorems on their own. It seems questionable whether they are preferable to **ZFC**, not to speak of a sufficiently developed dialetheist set theory.
- The arguments against a universal set partly [see the 2\textsuperscript{nd} category of objections] apply to **NF**-like systems, but not to dialetheist set theories having enough sets, even though many of them are inconsistent.
- The ontological alternatives to both a universal set and a theory of proper classes may be worth considering.
- The crucial point is how important Naïve Comprehension is. As Naïve Comprehension means trouble only with infinite sets this again may be rather a problem of the infinite [cf. Chap 12]. In the infinite **ZFC**-Comprehension doesn’t seem to bad.
- Ontological scruples to the side, everything would be fine if we can have a dialetheist set theory which at the same time is as easily and efficiently usable as **ZFC** is. More work has to be done in this area!
Giving Up Convention (T)?

- The main reason for dialetheism given in chapter 2 was the need for semantic universality. This universality requires defining semantic concepts within the language under investigation itself. The antinomies then ensue, most famously that of the Liar, involving the concept of truth.

- Formalizing the semantic antinomies, especially the Liar, makes use of the Tarskian idea of defining truth and Convention (T). Convention (T) is taken as a universal disquotational or correspondence fixing device:
  
  \[(T) \quad "p" \text{ is true (in } L) \text{ if and only if } p.\]

  One may think now that one can avoid some of the antinomies by giving up Convention (T).

- Giving up Convention (T) runs against its intuitive plausibility. So a compromise might be to restrict Convention (T) to some – consistent – contexts. (Remember that ZFC solved the Russell Paradox by restricting the naive/unrestricted comprehension axiom.)

- We take the biconditional in (T) to be mutual entailment ("$\leftrightarrow$") to have detachment. [cf. Chap. 16]
Convention (T) as a Default

- We can distinguish in Convention (T) the two directions
  
  $(T\rightarrow) \quad T("A") \rightarrow A \quad [or\ with\ "T"\ as\ operator: \ T(A) \rightarrow A]$
  
  $(T\leftarrow) \quad A \rightarrow T("A") \quad [or\ with\ "T"\ as\ operator: \ A \rightarrow T(A)]$

- Antinomies ensue only if (T) is employed to statements with semantic vocabulary (like "( ) is true", "( ) donates ( )"", "( ) means ( )"...). So we can start with:
  
  $(T1) \quad If\ A\ is\ a\ statement\ not\ containing\ semantic\ vocabulary:\ T("A") \leftrightarrow A.$

For statements about extra-lingual reality (T) holds, and this may be the reason for its intuitive force, since these are the statements we mostly care for.

- The problems with (T) concerns statements containing semantic vocabulary, but not all of them. We can even state:
  
  $(T2) \quad If\ A\ is\ not\ a\ statement\ referring\ to\ another\ statement\ containing\ semantic\ vocabulary:\ T("A") \leftrightarrow A$

In this case A refers to some non-semantic statement B, and A is true, according to (T), if B has the semantic property A says B has.
Convention (T) as a Default (II)

- We have problems if a semantic statement $A$ refers to another semantic statement $B$. In case of Liar chains $A$ and $B$ can be different statements:
  
  (1) Statement (2) is false.
  (2) Statement (1) is true.

  But not even all these cases lead to antinomies:
  
  (3) Statement (4) is false.
  (4) Statement (3) is false.

  lead to no antinomy, since (3) may be true, whereas (4) is false. The Truth-Teller is another case of harmless semantic circularity:

  (5) Statement (5) is true.

  The Truth-Teller may be taken either as true or as false, it certainly leads to no antinomy. So we can state as a fact:

  (F) Only in some cases of semantic statements referring to other semantic statements this leads to antinomies.

- Given that only sometimes problems ensue we might see (T) as a rule with minor exceptions, as a default.
Convention (T) as a Default (III)

- We can state the default:
  \[(T3) \text{ If A is a semantic statement referring to a semantic statement B, by default we have } T("A") \leftrightarrow A.\]
  one may also say "as a rule we have...".
- Default rules are nothing extraordinary. Within science and everyday life we often use them. Consider: "All birds can fly", "At the main station you get the connection", "Lessons are on Tuesday" ... In all these cases we don't give up the generalization once we hear of a penguin, miss a bus on a day with very bad weather or don't have lesson on a national holiday. We treat these cases as understandable exceptions to the otherwise true rule/generalization. There even are logics for this phenomenon: Default Logic or Non-Monotonic Logic!
- The problematic direction is \((T\leftarrow)\). One can argue \textit{ad hominem} against the default idea: Even if \((T\leftarrow)\) is not \textit{true} [i.e. \(T("A \rightarrow T("A")")\) is false], just \textit{as you say} it may be that \(A \rightarrow T("A")\), such that still antinomies ensue (Priest 1987, p.21). On the other hand the exceptions to (T) are taken to be antinomic contexts, default theory \textit{itself} isn't one.
Assessment of Default Convention (T)

- Restricting Convention (T) may block some of the semantic antinomies. Nevertheless there are three main problems with this approach:
  
  (i) The measure seems *ad hoc* in that we have to see the antinomy first to declare then that this was not the right context to apply Convention (T). On the other hand Default Logic is general framework, i.e. it was not invented to deal with the antinomies. Taking (T) as a default is so only an application of this wider approach. In concrete case, however, it is still *ad hoc* and not algorithmic.
  
  (ii) The approach was not fully formalized here. The problem with supposed solutions to antinomies is often that once we have formalized them we have a semantic machinery in place that allows to introduce new antinomies for the old (as with the Strengthened Liar in many-valued approaches).
  
  (iii) If there are exceptions to Convention (T) we seem to have statements with absent/restricted semantic properties. How is that possible?
Other Naïve Theories of Truth

- If the Naïve Theory of Truth is understood just as that theory that provides all the (T)-theorems for a language, there may be non-paraconsistent Naïve Theories of Truth that circumvent Tarski’s Theorem by avoiding some the ingredients of its proof.
- Typically this may be done by limiting expressive power, by resorting to some new version of a hierarchy (of truth predicates or evaluations) or by inventing special purpose conditionals for (T)-theorems.
- The problem with all this is, of course, that the main concern of dialetheism is not a Naïve Theory of Truth as narrowly understood as just done, but semantic closure. Approaches of the kind mentioned forsake in one way or the other the model that one and the same – more or the less intuitive – language/logic specifies its own semantics completely and without MYSTERY.

[Theories of the kind mentioned are, for example: (Field 2003), (Yablo 2003).]
Suszko's Thesis

- In Chap. 2 we have already seen that 3-valued approaches provide no alternative solution to the Liar paradox, because one introduce some version of Strengthened Liar.
- This result can be generalized. Suszko's Thesis is the claim that any many valued logic can be restated as a bivalent logic. This is done by providing a translation that maps all the designated values to 1 and the not designated truth values to 0. [cf. (Caleiro et al. 2005)].
- If the new – homomorphic – system allows for some antinomy so does the metalogical framework of the original many valued logic.
- Additionally note that this also hold for "pragmatic" 3-valued logics; the act of "supposing" a sentence to be true and to retract it if it turns out to be antinomic (an idea sometimes proposed as a pragmatic solution to the antinomies) rather resembles the adaptive logic procedure!
Schemata

- The semantic hierarchy failed in talking about the hierarchy itself and the general way to define semantic concepts.
- If one looks at the definitions repeated at ever higher levels, these definitions are very similar. They only differ in their indexes of levels. One might think, therefore, that one can give the *schematic form* of such a definition. That schema then would *express* what it is to define a semantic concept. Why shouldn’t that be enough?
- This leads to a difficult question: What is it to understand a schema? Compare: $A \supset (B \supset A) \text{ with } p \supset (q \supset p)$
The latter is an instance of the former. How do we take the former?
- If understanding the schema means "we know that *all* instances are valid" we are back at universal quantification.
- Quantifying over indexes that are only implicit – in case of a meta-semantic schema – does not make quantification about indexes and the resulting contradictions go away.
Schemata (II)

- The problem is old: Hilbert’s finitism was founded on the usage of schemata in contrast to generalizations. Using a schema was taken by Hilbert to mean "given any instance we accept it". So a schema talks about arbitrary instances, not about the totality of its instances.
- It certainly is no solution to take our ability to understand schemata as inexpressible. That would bring us back to some mystery.
- Furthermore one might doubt whether the understanding of any presented instance of a restricted definition is enough to understand the universal semantic concepts.
- Whether there is anything to this distinction between "any" and "all" seems to be an open question to me.
- For the moment I take it that the use of schemata is no alternative to non-hierarchical semantic concepts (and so to paraconsistency).
Science Fiction

- The ultimate alternative combining semantic closure with avoiding dialetheism is turning to (philosophical) science fiction.
- One may hint that the antinomies just show that there is something wrong with specific statements but not with semantic closure in general, so that one has to proceed with one's universal semantics while ignoring the antinomies, leaving them as oddities to the side.
- Now, the problem with this attitude is that once we start to make this procedure explicit and employ formalizations we have no way to avoid the antinomies (remember the Framework Hypothesis of chapter 2!).
- Here philosophical science fiction may intervene: it is the claim that although our current techniques of formal languages do not allow for an antinomy ignorance stance future developments will take care. We calmly have to refrain from the panic of inexpressible hierarchies as well as from the pessimism of dialetheism.
- No-one will complain about consistently closed semantics, I suppose. Given the well-known meta-logical theorems, Church's Thesis and the arguments of preceding chapters, the burden of proof – obviously – lays with the logical futurist.
Questions

- **(Q1)** Why can't the discrimination that we have an exception case to Convention (T) as a default rule not be algorithmic?

- **(Q2)** Lesniewski's argument against the Russell Set proceeds by definitions and axioms. (1) looks like set theoretical Comprehension. (2) merely says that classes are objects. (4), (5), (6), (8) are just consequences. So where is the decisive assumption? (3), (7)?

- **(Q3)** Why does Specker’s Theorem show (in NFU) that there are atoms, and so many of them?
Exercises

- (Ex1) Take the Liar and show that (T←) is the more problematic direction of Convention (T).

- (Ex2) Starting with the formalizations of (1) – (3) give a formalization of Lesniewski's argument.

- (Ex3) Proof that Naïve Comprehension and ZFC-Comprehension are equivalent for finite sets if there is the set of finite sets. How may one define the set of finite sets? Show that this set exists in ZFC.
Further Reading


On Hilbert's finitism and the use of schemata see Lavine's *Understanding the Infinite* (see reference Chap.11).

Frode Bjordal in "Truth in Perspective" (*International Philosophical Preprint Exchange*, 1996) takes (T←) as the culprit, but needs a modal logic with specially designed accessibility relations for his results.

Further Reading (II)